# STATISTICAL PROPERTIES OF NON-STATIONARY DYNAMICAL SYSTEMS WITH INTERMITTENCY

JUHO LEPPÄNEN

Academic dissertation

*To be presented for public examination*
*with the permission of the Faculty of Science of the University of Helsinki*
*in Auditorium CK112, Exactum, on 14 August 2018 at 12.00*

Department of Mathematics and Statistics
Faculty of Science
University of Helsinki
2018

**Supervisor**
Mikko Stenlund, University of Helsinki, Finland

**Pre-examiners**
Tommi Sottinen, University of Vaasa, Finland
Matti Vihola, University of Jyväskylä, Finland

**Opponent**
Tomas Persson, Lund University, Sweden

**Custos**
Antti Kupiainen, University of Helsinki, Finland

ISBN 978-951-51-4421-8 (nid.)
ISBN 978-951-51-4422-5 (PDF)

Unigrafia,
Helsinki 2018

# Statistical properties of non-stationary dynamical systems with intermittency

Juho Leppänen

Department of Mathematics and Statistics
P.O. Box 68 (Gustaf Hallströmin katu 2b)
FI-00014 University of Helsinki, Finland
juho.leppanen@helsinki.fi

**Abstract**

The problems addressed in this thesis revolve around two types of non-stationary dynamical systems: sequential compositions of interval maps with a neutral fixed point (Pomeau-Manneville maps) and intermittent quasistatic systems. Both systems are non-uniformly expanding and time-dependent, and (typically) lack invariant measures. The evolution of states under a sequential system is described by a sequence of varying self-maps $T_1, T_2, \dots$ of a phase space $X$. Such constructions are motivated by applications to non-equilibrium processes in nature, where the map $T_n$ describing how a state evolves from time $n$ to $n+1$ should depend on $n$. Quasistatic systems on the other hand draw inspiration from thermodynamics and model situations where the observed system transforms (infinitesimally) slowly with time due to external influence. At any given time the system is at an equilibrium, but over a long time span the equilibrium slowly changes. The thesis consists of an introductory part and three research articles. The first and third article are about quasistatic systems, while the second article deals with sequential systems.

The second article is motivated by multivariate normal approximation for Pomeau-Manneville maps. The main result is a functional correlation bound widely useful for showing limit theorems in the sequential settting. We prove the result by modifying a technique of Liverani, Saussol, and Vaienti, which is based on a probabilistic approximation of the deterministic system. We present two applications of the result for a single Pomeau-Manneville map, by showing that the bound implies the correlation-decay conditions of the normal approximation methods due to Pène-Rio and Stein. Both methods yield a multivariate central limit theorem with an estimate on the rate of convergence. The rate produced by the former method is optimal with respect to the Kantorovich (or Wasserstein) metric. The latter method is suitable also for normal approximation in non-stationary settings.

In the first article we introduce the intermittent quasistatic system and obtain several tools for its further analysis, including $L^1$-perturbation estimates for the transfer operators. The main result is an almost sure ergodic theorem for the time-averages of the model. The proof, which is partly based on a general theory developed by Stenlund, makes extensive use of the polynomial memory loss bound shown recently by Aimino et al. The third article builds on the results of the first two articles. By solving a well-posed martingale problem, we show that limiting distributional behavior of intermittent quasistatic systems can be characterized by a stochastic diffusion process. The result extends that shown by Dobbs and Stenlund for a class of uniformly expanding quasistatic systems.

## Acknowledgements

First and foremost, I thank my supervisor Mikko Stenlund for his guidance and support, for his generous advice, and for many discussions. Mikko showed me into ergodic theory and introduced me to the topics of this thesis. Many times I have relied on his insight to develop ideas and progress in my studies. Since the day we met I have been deeply impressed by Mikko's attitude toward mathematics and research.

Secondly, I thank Antti Kupiainen for the support and encouragement I have received over these years, and for acting as the custos of the thesis defense. I am honored for having Tomas Persson as my opponent during the defense. I thank the pre-examiners of the thesis, Tommi Sottinen and Matti Vihola, for carefully reviewing the manuscript and allowing me to graduate on a tight schedule, and for all their valuable comments.

I am indebted to Romain Aimino for the discussions we had in 2016, and especially for the detailed explanation he gave regarding the proof for the positivity of the kernel function[1]. I convey my sincere gratitude to Carlangelo Liverani for all the insightful comments he gave me this summer. They considerably strengthened my understanding of some of the results in the thesis.

I thank Olli Hella for being a good colleague and friend. I appreciate there being a fellow student at the department with whom I could share my views about mathematics and doctoral student life in general. Our mathematical discussions sometimes led to new discoveries and often taught me to be more critical toward my work.

I have been studying at the Helsinki Department of Mathematics and Statistics for almost eight years. Highlights of this time were the numerous inspiring courses I attended. I thank my teachers, who excited me with their enthusiasm for mathematics and set me on a path toward rigor and devotion.

I thank my family and friends outside academia for their support, and for occasionally offering me an escape from mathematics. I use the opportunity to thank Topi for more than two decades of continued friendship.

The funding I received from the Jane and Aatos Erkko Foundation, the Emil Aaltosen Säätiö, and the Doctoral Programme in Mathematics and Statistics (Domast) made this research possible and is gratefully acknowledged.

Lastly, a very special thanks to Sayuri.

Juho Leppänen
July 2018, Helsinki


[1] Theorem 1.3

## LIST OF INCLUDED ARTICLES

This thesis consists of an introductory part and three research articles. During the introduction we refer to these articles by their respective roman numerals.

**(I)** Juho Leppänen and Mikko Stenlund. Quasistatic dynamics with intermittency. *Math. Phys. Anal. Geom.*, 19(2):Art. 8, 23, 2016. arXiv:1510.02748

**(II)** Juho Leppänen. Functional correlation decay and multivariate normal approximation for nonuniformly expanding maps. *Nonlinearity*, 30(11):4239, 2017. arXiv:1702.00699

**(III)** Juho Leppänen. Intermittent quasistatic dynamical systems: weak convergence of fluctuations. *Nonauton. Dyn. Syst.*, 1:8-34, 2018. arXiv:1710.11371.

Article (I) is a joint work with my supervisor Mikko Stenlund, who introduced to me the problem and all the main ideas. The first three sections of the paper were devised and written entirely by Stenlund. They focus mainly on motivation and background, and contain little new research. The work for sections 4 through 6 was done together.

## CONTENTS

## 1. BACKGROUND

The topics of this thesis lie at the interface of non-stationary and non-uniformly hyperbolic dynamics. Non-uniformly hyperbolic systems have been studied extensively in the context of maps with neutral fixed points, at least since the Pomeau-Manneville map was proposed as a model for the intermittency of turbulent flows [17, 42]. We study the following version of the Pomeau-Manneville map introduced by Liverani, Saussol, and Vaienti in [32]: for each $\alpha \in (0,1)$, define the map $T_\alpha : [0,1] \to [0,1]$ by

$$T_\alpha(x) = \begin{cases} x(1+2^\alpha x^\alpha) & \forall x \in [0,1/2), \\ 2x-1 & \forall x \in [1/2,1]. \end{cases} \tag{1}$$

Such maps are often called "intermittent" because of their dynamical characteristics: the dynamics is strongly chaotic due to expansion, except near the neutral fixed point at the origin. Once the trajectory of a point lands near the origin, it stays there for a possibly large number of iterations, until expansion eventually has noticeable effect again and the trajectory returns to the strongly chaotic region. Local expansion around the origin weakens as $\alpha$ grows. On the other hand, as $\alpha \downarrow 0$ the neighborhood in which $T'_\alpha \approx 1$ becomes ever smaller, and at $\alpha = 0$ we arrive at the uniformly expanding angle doubling map.

It was shown by Pianigiani [41] in 1980 that expanding maps with finitely many neutral fixed points admit absolutely continuous invariant measures. The map $T_\alpha$ preserves an absolutely continuous probability measure, which we denote by $\hat{\mu}_\alpha$. By a result due to Thaler [51], the density $\hat{h}_\alpha$ of $\hat{\mu}_\alpha$ becomes ever more concentrated around the origin as $\alpha$ grows with a sharp estimate[2] $\hat{h}_\alpha(x) \sim x^{-\alpha}$. In fact, it follows from [32] that $\hat{h}_\alpha$ belongs to the convex cone

$$\begin{aligned} \mathcal{C}_*(\alpha) = \{ f \in C((0,1]) \cap L^1 \ : f \geq 0, f \text{ decreasing}, \\ x^{\alpha+1} f \text{ increasing}, f(x) \leq 2^\alpha (2+\alpha) x^{-\alpha} m(f) \}, \end{aligned}$$

where $m$ denotes the Lebesgue measure and $m(f) = \int f \, dm$.

Correlation decay for intermittent maps was studied in the 90's, first by Mori [36] and Lambert et al. [29] in linear settings. The more difficult non-linear case was then treated in [27, 32, 52]. Young's [52] highly general approach was based on an abstract "tower" model where the speed of correlation decay is determined by the tail of the return time function. The model covers many non-uniformly expanding systems, e.g. Viana maps [3] and certain unimodal maps [11], in addition to intermittent maps. For maps such as $T_\alpha$ that satisfy $-xT''_\alpha(x) \sim x^\alpha$ near the fixed point $x = 0$, the results of [52] show that the correlation functions have an upper bound of order $n^{1-1/\alpha}$. A lower bound of the same order was obtained by Hu [27] via transfer operator techniques. Utilizing a randomly perturbed version of the transfer operator, Liverani et al. [32] showed an upper bound of order $n^{1-1/\alpha}(\log n)^{1/\alpha}$ for the map (1). The rate is nearly optimal: it is the same as that of Hu's lower bound apart from the logarithmic correction. Later Sarig [46] and Gouëzel [21] obtained sharp correlation decay rates in the Young tower setting of [52].

In this thesis we focus on two particular classes of non-stationary systems, both of which arise as suitable (non-random) compositions of the intermittent maps (1). The first class is described by sequential compositions of the form $T_{\alpha_n} \circ \cdots \circ T_{\alpha_1}$ where

[2] We denote $g(x) \lesssim f(x)$ if there exists a system constant $C > 0$ such that $f(x) \leq Cg(x)$ for all $x$ in the domain of $f$ and $g$. If $C$ depends on some additional parameter $\alpha$, we indicate this by writing $g(x) \lesssim_\alpha f(x)$ instead. Then $f(x) \sim g(x)$ means $g(x) \lesssim f(x)$ and $f(x) \lesssim g(x)$.

each $T_{\alpha_n}$ is a map in the family (1). The second class is a related construction called intermittent quasistatic systems where the time-evolution is given by compositions $T_{\alpha_{n,k}} \circ \cdots \circ T_{\alpha_{n,1}}$ of maps whose parameters $\alpha_{n,k}$ belong to the the same "level" of a triangular array $\{\alpha_{n,k} \in [0,1) : 0 \le k \le n,\ n \in \mathbb{N}\}$. Sequential intermittent systems are discussed below and in Section 2, and quasistatic systems in Sections 1.2 and 3 of this introduction.

1.1. **Memory loss.** Non-stationary (or non-autonomous) systems model time-dependent dynamics and physical processes that take place in evolving environments. Important examples include Sinai billiards where some of the scatterers move due to bombardment by lightweight particles [50], and open systems with moving holes [35]. To facilitate the analysis of such systems, Ott, Stenlund and Young [39] considered in 2009 the notion of memory loss – a counterpart for the stationary notion of correlation decay. The system is said to lose its memory (in the statistical sense) if the time-evolutions $\mu_n, \nu_n$ of any two sufficiently regular initial distributions $\mu_0, \nu_0$ satisfy $\|\mu_n - \nu_n\| \to 0$ as $n \to \infty$ with respect to some suitable notion of distance $\| \cdot \|$. The condition can be interpreted as all regular distributions being attracted by the same moving target in the space of measures. We emphasize that, unlike in the case of a stationary system with good statistical properties, measures evolving under non-stationary dynamics typically fail to converge toward any invariant distribution due to the dynamics changing with time.

Of considerable importance is the rate at which memory is lost. In [39], Ott et al. showed by using an adaptation of the coupling method [34, 52] that time-dependent compositions $T_n \circ \cdots \circ T_1$ of uniformly expanding and one-dimensional piecewise expanding maps $T_n : X \to X$ satisfy a rapid rate in the strong sense

$$\int |\varphi_n - \psi_n| \, d\mu \lesssim_{\varphi,\psi} \theta^n, \quad (0 < \theta < 1) \tag{2}$$

where $\varphi_n$ and $\psi_n$ denote the time-evolutions of some initial densities $\varphi$ and $\psi$ with respect to a suitable reference measure $\mu$ on $X$. In this case it is said that the system loses its memory exponentially. Given an observable $f : X \to \mathbb{R}$, one may for instance use the bound to study limiting properties of the Birkhoff sums $\sum_{k=0}^{n-1} f \circ T_k \circ \cdots \circ T_1$. The particular problem of normal convergence has been considered in [8, 9, 13, 37].

The rate of memory loss was studied also in [24, 35, 48, 50] and in [1]. The first four papers again obtained exponential rates: [48] for topologically transitive Anosov diffeomorphisms, [50] for Sinai billiards with moving scatterers, [24] for piecewise expanding maps in higher dimension, and [35] for piecewise smooth open systems. The study [1] of Aimino et al. from 2015 established for the first time a sub-exponential rate of memory loss. The authors showed by building on the probabilistic approximation method of Liverani et al. [32] that sequential compositions $T_{\alpha_n} \circ \cdots \circ T_{\alpha_1}$ of maps in the intermittent family (1) satisfy a rate of memory loss that corresponds to the polynomial correlation decay rate of [32]. An extension of the method essentially yields the main result of article (II), which is why we next discuss the contents of [1] in more detail.

Let us fix a number $\beta_* \in (0,1)$ and denote $\mathcal{C}_* = \mathcal{C}_*(\beta_*)$. We call a sequence $(T_n)_{n \ge 1}$ of intermittent maps $T_n = T_{\alpha_n}$ admissible, if the parameters $\alpha_n$ satisfy $\alpha_n \le \beta_*$ for all $n \in \mathbb{N}$. We denote by $\mathcal{L}_\alpha : L^1(m) \to L^1(m)$ the transfer operator associated to $T_\alpha$; that is,

$$\mathcal{L}_\alpha h(x) = \sum_{y \in T_\alpha^{-1}\{x\}} \frac{h(y)}{T_\alpha'(y)}.$$

Given any admissible sequence $(T_n)_{n\geq 1}$ of maps we abbreviate $\mathcal{L}_n = \mathcal{L}_{\alpha_n}$.

By the following result from [1], concatenations of transfer operators decay polynomially in $L^1(m)$:

**Theorem 1.1.** *Suppose $(T_n)_{n\geq 1}$ is an admissible sequence of maps, and $\varphi, \psi \in \mathcal{C}_*$ with $m(\varphi) = m(\psi)$. Then, for all $n \in \mathbb{N}$,*

$$\|\mathcal{L}_n \cdots \mathcal{L}_1(\varphi - \psi)\|_{L^1(m)} \leq C(m(\varphi) + m(\psi))\rho(n),$$

*where $\rho(n) = n^{1-\frac{1}{\beta_*}}(\log n)^{\frac{1}{\beta_*}}$ for $n \geq 2$, and $\rho(0) = \rho(1) = 1$. The constant $C > 0$ in the bound depends only on the system $T_{\beta_*}$.*

*Remark* 1.2. In fact, the authors of [1] considered a slightly modified version of the map (1) but pointed out that their approach works equally well for other related maps, including (1).

A central idea in the strategy of [1, 32] was to perform a random perturbation in order to suppress the intermittency effect. The idea is made rigorous by the notion of a perturbed transfer operator, which in the time-dependent case is defined by

$$\mathbb{L}_{\varepsilon,m} = \mathcal{L}_{m+n_\varepsilon-1} \cdots \mathcal{L}_m \mathbb{A}_\varepsilon, \quad m \geq 1,$$

where $\mathbb{A}_\varepsilon$ is the averaging operator[3]

$$\mathbb{A}_\varepsilon f(x) = \frac{1}{2\varepsilon}\int_{B_\varepsilon(x)} f\,dm,$$

and $n_\varepsilon \sim \varepsilon^{-\beta_*}$ is a positive integer given by the following important result from [1]:

**Theorem 1.3.** *There exist $\omega \in (0,1)$ and for all $\varepsilon > 0$ a number $n_\varepsilon \sim \varepsilon^{-\beta_*}$ such that the kernel function*

$$K_{\varepsilon,m}(x,z) = \frac{1}{2\varepsilon}\mathcal{L}_{n_\varepsilon+m-1} \cdots \mathcal{L}_m \mathbf{1}_{B_\varepsilon(z)}(x)$$

*satisfies*

$$K_{\varepsilon,m}(x,z) \geq \omega \quad \forall x,z \in [0,1],\ \forall m \in \mathbb{N}. \tag{3}$$

Proving (3) was the primary obstacle to obtaining Theorem 1.1. To demonstrate the usefulness of the lower bound, let us write

$$\mathbb{L}_{\varepsilon,m} f(x) = \int_0^1 K_{\varepsilon,m}(x,z)\cdot f(z)\,dz,$$

Then, let $g \in L^1(m)$ with $m(g) = 0$, and set $I^+_{\varepsilon,m} = (\mathbb{L}_{\varepsilon,m} g)^{-1}[0,\infty)$ and $I^+ = g^{-1}[0,\infty)$. Since $0 = m(g) = m(\mathbb{L}_{\varepsilon,m} g)$,

$$\begin{aligned}\int_0^1 |\mathbb{L}_{\varepsilon,m} g|\,dm &= 2\int_{I^+_{\varepsilon,m}}\int_0^1 K_\varepsilon(x,z)g(z)\,dz\,dx = 2\int_{I^+_{\varepsilon,m}}\int_0^1 (K_\varepsilon(x,z)-\omega)g(z)\,dz\,dx \\ &\leq 2\int_0^1\int_{I^+}(K_\varepsilon(x,z)-\omega)g(z)\,dz\,dx = (1-\omega)\|g\|_{L^1(m)}.\end{aligned}$$

Iterating the estimate then yields

$$\|\mathbb{L}_{\varepsilon,(k-1)n_\varepsilon+m} \cdots \mathbb{L}_{\varepsilon,m} g\|_1 \leq (1-\omega)^k \|g\|_{L^1(m)} \quad \forall k \in \mathbb{N}.$$

In other words, (3) implies that the perturbed transfer operator decays at an exponential rate. Theorem 1.1 essentially follows from this, after observing that for functions

[3] We denote $B_\varepsilon(x) = \{y \in [0,1) : d(x,y) \leq \varepsilon\}$, where $d(x,y) = \min\{|x-y|, 1-|x-y|\}$.

$\varphi \in \mathcal{C}_*$, the $L^1$-distance between $\mathbb{L}_{\varepsilon,m}\varphi$ and $\mathcal{L}_{m+n_\varepsilon-1}\cdots\mathcal{L}_m\varphi$ decays polynomially in $\varepsilon$, in fact at the rate $\varepsilon^{1-\beta_*}$.

Theorem 1.1 implies the polynomial decay of (non-stationary) correlations. It was shown in [1] that for all $f \in C^1([0,1])$ and $g \in C^\infty([0,1])$,

$$\left| \int f \cdot g \circ \widetilde{T}_n \, d\mu - \int f \, d\mu \int g \circ \widetilde{T}_n \, d\mu \right| \lesssim_f \|g\|_\infty \rho(n), \tag{4}$$

where $\widetilde{T}_n = T_n \circ \cdots \circ T_1$. Note that in the case $\mu = m$ and $f \in \mathcal{C}_*$ the bound is an immediate corollary of Theorem 1.1. The full result is obtained from this by observing that products $fh$ of functions $f \in C^1([0,1])$ and $h \in \mathcal{C}_*$ can be written as differences $g_1 - g_2$ of suitable cone functions. We comment more on this in Section 3.1.

1.2. **Quasistatic systems.** Articles (I) and (III) deal with quasistatic systems. In thermodynamics, the term quasistatic refers to an idealized process where the observed system transforms infinitesimally slowly due to external influences. Motivated by such setups, Dobbs and Stenlund [14] proposed in 2015 the following class of deterministic non-stationary systems.

**Definition 1.4** (Discrete time QDS). *Let $(X, \mathscr{F})$ be a measurable space, $\mathcal{M}$ a topological space whose elements are measurable self-maps $T : X \to X$, and $\mathbf{T}$ a triangular array of the form*

$$\mathbf{T} = \{T_{n,k} \in \mathcal{M} \; : \; 0 \le k \le n, \; n \ge 1\}.$$

*If there exists a piecewise continuous curve $\tau : [0,1] \to \mathcal{M}$ such that*[4]

$$\lim_{n\to\infty} T_{n,\lfloor nt \rfloor} = \tau_t \tag{5}$$

*for all $t$, we say that $(\mathbf{T}, \tau)$ is a* quasistatic dynamical system (QDS) *with* state space $X$ *and* system space $\mathcal{M}$.

The limit curve $\tau$ models the evolution of a slowly transforming system. Regularity properties of $\tau$ and the rate of convergence in (5) typically play an important role in the analysis of a particular QDS. Usually there does not exist a measure invariant for all $\tau_t$, but there still exists a family of measures $\{\hat{\nu}_t\}_{t\in[0,1]}$ with some nice properties, such that each $\hat{\nu}_t$ is invariant for $\tau_t$.

The dynamics of the QDS $(\mathbf{T}, \tau)$ are described by the triangular array $\mathbf{T}$: given an initial state $x \in X$, $x_{n,k} = T_{n,k} \circ \cdots \circ T_{n,1}(x)$ is the state of the system after $k$ steps on the $n$th level of the array. Since $T_{n,1} \approx \tau_0$ typically differs considerably from $T_{n,n} \approx \tau_1$ for large $n$, it is not possible to describe statistical behavior by directly studying limits of $T_{n,k}$ for fixed $k$. Rather it is necessary to analyze the entire curve $t \mapsto T_{n,\lfloor nt \rfloor}$ and see how it behaves in the limit. For this purpose we define for a given measurable function $f : X \to \mathbb{R}$ the functions $S_n : X \times [0,1] \to \mathbb{R}$ by

$$S_n(x,t) = \int_0^{nt} f(x_{n,\lfloor s \rfloor}) \, ds.$$

In other words, $S_n(x,t)$ is a piecewise linear interpolation of the Birkhoff type sum $\sum_{k=0}^{\lfloor nt \rfloor - 1} f(x_{n,k})$. Note that, given an initial distribution $\mu$ for $x \in X$, $x \mapsto S_n(x,\cdot)$ can be viewed as a random element with values in the space $C^0([0,1])$ of continuous real-valued functions. Then it is natural to ask whether $\zeta_n(x,t) = n^{-1}S_n(x,t)$ converges in some sense. In the degenerate case $T_{n,k} = T_{1,0}$ for all $0 \le k \le n$, Birkhoff's theorem

[4] For any real number $s \ge 0$, $\lfloor s \rfloor$ denotes the integer part of $s$.

guarantees that $\zeta_n(x,t) \to t\int_X f\,d\mu$ almost surely, provided that $\mu$ is $T_{1,0}$-invariant and $T_{1,0}$ is ergodic. Below we state an ergodic theorem from [49] for a more general class of QDSs.

For all $0 \le k \le n$, we denote

$$f_{n,k} = f \circ T_{n,k} \circ \cdots \circ T_{n,1},$$

and

$$\bar{f}_{n,k} = f_{n,k} - \mu(f_{n,k}).$$

Then, we set

$$\bar{\zeta}_n(x,t) = \int_0^t \bar{f}_{n,\lfloor ns\rfloor}(x)\,ds,$$

and

$$\zeta(t) = \int_0^t \hat{v}_s(f)\,ds.$$

**Theorem 1.5.** *Let $f : X \to \mathbb{R}$ be a bounded measurable function and $\mu$ a probability measure. Suppose the following conditions hold:*

(i) *The map $t \mapsto \hat{v}_t(f)$ is measurable;*
(ii) $\lim_{n\to\infty}\mu(f_{n,\lfloor nt\rfloor}) = \hat{v}_t(f)$ *for almost every $t \in [0,1]$;*
(iii) *For all integers $2 \le l \le 4$, $j \in \{1, l-1\}$ and $0 \le k_1 \le \cdots \le k_l$,*

$$|\mu(f_{n,k_1}\cdots f_{n,k_l}) - \mu(f_{n,k_1}\cdots f_{n,k_j})\mu(f_{n,k_{j+1}}\cdots f_{n,k_l})| \lesssim_f \Phi(k_{j+1}-k_j),$$

*where $\Phi(s) = s^{-1}(\log s)^{-2}$ if $s \ge 2$, and $\Phi(s) = 2^{-1}(\log 2)^{-2}$ if $0 < s < 2$.*

*Then,*

$$\lim_{n\to\infty}\sup_{t\in[0,1]}|\zeta_n(x,t) - \zeta(t)| = 0 \tag{6}$$

*for almost every $x$ with respect to $\mu$.*

Condition (iii) presumes a sufficiently rapid polynomial decay of correlations and, by Lemma 5.1 in [49], implies the fourth moment condition

$$\sum_{n=1}^{\infty}\mu(|\bar{\zeta}_n(\cdot,t)|^4) < \infty \quad \forall t \in [0,1]. \tag{7}$$

In particular, it follows that $\lim_{n\to\infty}\bar{\zeta}_n(x,t) = 0$ holds almost surely for any fixed $t \in [0,1]$. Since $f$ is bounded, the family of functions $t \mapsto \bar{\zeta}_n(x,t)$ is uniformly Lipschitz continuous, and this suffices for obtaining (see Lemma 4.1 of [49])

$$\lim_{n\to\infty}\sup_{t\in[0,1]}|\bar{\zeta}_n(x,t)| = 0$$

almost surely with respect to $\mu$. Proving (6) under condition (iii) in this way reduces to showing

$$\lim_{n\to\infty}\sup_{t\in[0,1]}\left|\int_0^t \mu(f_{n,\lfloor ns\rfloor}) - \hat{v}_s(f)\,ds\right| = 0,$$

which is a direct consequence of condition (ii). Note that (5) and the piecewise continuity of $\tau$ suggest[5]

$$\begin{aligned}\mu(f_{n,\lfloor ns\rfloor}) &= (T_{n,\lfloor ns\rfloor}\circ\cdots\circ T_{n,\lfloor n(s-\varepsilon)\rfloor+1})_*(T_{n,\lfloor n(s-\varepsilon)\rfloor}\circ\cdots\circ T_{n,1})_*\mu(f)\\ &\approx (\tau_s)_*^{\lfloor n\varepsilon\rfloor}(T_{n,\lfloor n(s-\varepsilon)\rfloor}\circ\cdots\circ T_{n,1})_*\mu(f),\end{aligned}$$

whenever $\varepsilon > 0$ is small and $n$ is large. Thus, we can expect condition (ii) if the pushforward-measures $(T_{n,\lfloor n(s-\varepsilon)\rfloor}\circ\cdots\circ T_{n,1})_*\mu$ belong to a class of measures $\nu$ such that $\lim_{n\to\infty}(\tau_s)_*^n\nu = \hat{\nu}_s$. The latter convergence corresponds to a memory-loss property of the autonomous system $\tau_s$.

Conditions (i)-(iii) of Theorem 1.5 were verified in [49] for quasistatic billiards and quasistatic expanding systems. The former model obtains in the limit of a dispersing billiard with infinitesimally slowly moving scatterers. For the latter model we state the rigorous definition from [49].

(M1) The system space $\mathcal{M}$ consists of all $C^2$ expanding maps $T : S^1 \to S^1$ of the circle $S^1 = \mathbb{R}/\mathbb{Z}$ that satisfy $\inf T' \geq \lambda$ and $\|T''\|_\infty = \sup_{x\in S^1}|T''x| \leq A_*$ for the same $\lambda > 1$ and $A_* > 0$. The space $\mathcal{M}$ is endowed with the metric

$$d_{C^1}(T_1, T_2) = \sup_{x\in S^1} d(T_1x, T_2x) + \|T_1' - T_2'\|_\infty,$$

where $d$ denotes the natural metric on $S^1$.

(M2) The curve $\tau : [0,1] \to \mathcal{M}$ is Hölder continuous of order $\eta \in (0,1)$, such that

$$\sup_{0\leq t\leq 1} d_{C^1}(T_{n,\lfloor nt\rfloor}, \tau_t) \lesssim n^{-\eta}.$$

The rate of convergence in (M2) is the natural one indicated by the "equipartition" $T_{n,k} = \tau_{kn^{-1}}$, although in general the maps $T_{n,k}$ are not required to lie in the range of $\tau$. Each $\tau_t$ has a unique absolutely continuous invariant probability measure $\hat{\nu}_t$ whose density lies in $\mathrm{Lip}(S^1) = \{h : h : S^1 \to \mathbb{R} \text{ Lipschitz continuous}\}$. We equip the space $\mathrm{Lip}(S^1)$ with the norm $\|h\|_{\mathrm{Lip}} = \mathrm{Lip}(h) + \|h\|_\infty$, where

$$\mathrm{Lip}(h) = \sup_{x\neq y}\frac{|h(x) - h(y)|}{d(x,y)}.$$

The study [14] concerned further statistical properties of the quasistatic expanding system. Let $\mu$ be an initial distribution of $x \in S^1$, and for all $t \in [0,1]$ denote $\hat{f}_t = f - \hat{\nu}_t(f)$. Instead of the mean $\zeta_n(x,t)$, the authors of [14] studied fluctuations at a finer scale, by looking at

$$\xi_n(x,t) = n^{-\frac{1}{2}}S_n(x,t) - n^{-\frac{1}{2}}\mu(S_n(\cdot,t)),$$

where $S_n$ has been redefined (in the obvious way) for the quasistatic expanding system. Here we often hide the $x$-dependence and denote $\xi_n(t) = \xi_n(x,t)$. For each $n \in \mathbb{N}$, the map $x \mapsto \xi_n(x,\cdot)$ is a random element with values in $C^0([0,1])$, and we denote its distribution (with respect to $\mu$) by $\mathbb{P}_n^\mu$. The weak limit of $(\mathbb{P}_n^\mu)$ was identified in [14] as the law of a diffusion process:

[5]The pushforward $(T)_*\mu$ of a measure $\mu$ by a measurable map $T : X \to X$ is defined by $(T)_*\mu(A) = \mu(T^{-1}A)$ for all $A \in \mathscr{F}$.

**Theorem 1.6.** *Suppose the observable $f : S^1 \to \mathbb{R}$ is Lipschitz continuous and $\mu$ is absolutely continuous with Lipschitz continuous density. Then the measures $\mathbb{P}_n^\mu$ converge weakly to the law of the processs*

$$\xi(t) = \int_0^t \hat{\sigma}_s(f)\, dW_s,$$

*where $W$ denotes a standard Brownian motion, the stochastic integral is defined in the sense of Itō, and*

$$\hat{\sigma}_t^2(f) = \lim_{m\to\infty} \hat{\nu}_t \left[ \left( \frac{1}{\sqrt{m}} \sum_{k=0}^{m-1} \hat{f}_t \circ \tau_t^k \right)^2 \right].$$

*Remark* 1.7. We have assumed for simplicity that the centering measure $\mu$ of $\xi_n$ coincides with the initial measure, but the conclusion remains true for a wide class of centerings other than $\mu(S_n(\cdot, t))$. After changing the centering, regularity assumptions on the density of the initial measure can be dropped. A proof for these generalizations would proceed with applications of the Portmanteau theorem and a density argument.

The proof of the theorem was based on solving a martingale problem that corresponds to the expression of $\xi$. Tightness of the measures $\mathbb{P}_n^\mu$ was established first by invoking the Kolmogorov-Chentsov criterion, after observing that

$$\mu[[\xi_n(t+\delta) - \xi_n(t)]^4] \lesssim \|f\|_{\mathrm{Lip}}^4 \delta^2. \tag{8}$$

The bound is a consequence of exponential correlation decay. Tightness then guarantees the existence of a weakly convergent subsequence $(\mathbb{P}_{n_k}^\mu)_{k\ge 1}$. To identify the weak limit $\mathbb{P} = \lim_k \mathbb{P}_{n_k}^\mu$, the following Dynkin formula was shown: for all $t \in [0,1]$ let $\pi_t : C^0([0,1]) \to \mathbb{R}$ be the evaluation functional $\pi_t(\omega) = \omega(t)$. Then,

$$\mathbb{E}[A \circ \pi_t] = \mathbb{E}[A \circ \pi_0] + \frac{1}{2} \int_0^t \mathbb{E}[A'' \circ \pi_s] \hat{\sigma}_s^2(f)\, ds \quad \forall A \in C_c^\infty(\mathbb{R}), \tag{9}$$

where $\mathbb{E}$ denotes expectation with respect to $\mathbb{P}$, and $C_c^\infty(\mathbb{R})$ denotes the collection of functions in $C^\infty(\mathbb{R})$ with compact support. The formula leads to the conjecture that the sought limit process $\xi$ is a diffusion with

$$\mathscr{L}_t = \frac{1}{2} \hat{\sigma}_t^2(f) \frac{d^2}{dx^2}$$

as its generator. This means that $\xi$ should solve the stochastic differential equation $d\xi(t) = \hat{\sigma}_t(f)\, dW_t$. A strong solution to the latter equation always exists, and its unique law $Q$ is characterized by the following martingale property [45]:

**Proposition 1.8.** *The measure $Q$ is the unique measure such that $Q(\pi_0 = 0) = 1$ and for all $A \in C_c^\infty(\mathbb{R})$ the process*

$$M_t = A \circ \pi_t - A \circ \pi_0 - \int_0^t \mathscr{L}_s A \circ \pi_s\, ds, \quad t \in [0,1],$$

*is a martingale with respect to $Q$ and the filtration $(\mathfrak{F}_t)_{t\in[0,1]}$, where $\mathfrak{F}_t$ is the sigma-algebra on $C^0([0,1])$ generated by $\{\pi_s : 0 \le s \le t\}$.*

Given the foregoing characterization we see that, to obtain Theorem 1.6, it suffices to show that $\mathbb{P}$ satisfies the above martingale property, for then we must have $\mathbb{P} = Q$.

The proof of the martingale property was based on certain auxiliary results that granted control over the second moments $\mu[[\xi_n(t) - \xi_n(s)]^2]$. Key ingredients for controlling the second moments were exponential memory loss and perturbation estimates for the transfer operators $\mathcal{L}_T$ of the expanding maps $T \in \mathcal{M}$. In [14] it was shown that for any Lipschitz continuous function[6] $g : S^1 \to \mathbb{R}$ with $m(g) = 0$,

$$\|\mathcal{L}_{T_n} \cdots \mathcal{L}_{T_1} g\|_{L^1(m)} \lesssim_g \theta^n, \tag{10}$$

where $\mathcal{L}_{T_n} \cdots \mathcal{L}_{T_1}$ is any $n$-concatenation of transfer operators. Moreover, if the transfer operators are viewed as mappings $\mathrm{Lip}(S^1) \to C^0(S^1)$, then

$$\|\mathcal{L}_{T_1} - \mathcal{L}_{T_2}\|_{\mathrm{Lip}\to C^0} \lesssim d_{C^1}(T_1, T_2), \tag{11}$$

where $\| \cdot \|_{\mathrm{Lip}\to C^0}$ denotes the operator norm. Bound (10) is a strong type of exponential memory loss that was established via coupling. In the proof of the perturbation estimate (11) it was important that the transfer operators $\mathcal{L}_T$ map from a space of regular functions to a space of less regular functions. The two bounds can be used, for instance, to show condition (ii) of Theorem 1.5.

## 2. Normal approximation for intermittent maps

We work in the setting of the intermittent family (1); recall that $\hat{\mu}_\alpha$ denotes the absolutely continuous invariant probability measure associated to $T_\alpha$. Then suppose $f : [0, 1] \to \mathbb{R}$ is a Lipschitz continuous function with $\hat{\mu}_\alpha(f) = 0$. By Hu's result [27], we know that $\hat{\mu}_\alpha(f \cdot f \circ T_\alpha^n) = O(n^{1-1/\alpha})$ so that if $\alpha < 1/2$, the sum $\sum_{n=0}^{\infty} \hat{\mu}_\alpha(f \cdot f \circ T_\alpha^n)$ converges absolutely. It follows immediately by a general result due to Liverani [31] that the central limit theorem (CLT) holds, but in fact much more is known. Gouëzel [22] showed that if $\alpha < 1/3$ and $f$ can not be written as $g - g \circ T_\alpha$, then the Berry-Esseen theorem holds. That is,

$$\left| \hat{\mu}_\alpha \left( \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} f \circ T_\alpha^k \le x \right) - \int_{-\infty}^{x} \phi_{\sigma^2}(y)\, dy \right| \lesssim_f n^{-\frac{1}{2}},$$

where $\phi_{\sigma^2}$ denotes the density function of the centered normal distribution $\mathcal{N}(0, \sigma^2)$ with positive variance

$$\sigma^2 = \lim_{n\to\infty} \hat{\mu}_\alpha \left[ \left( \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} f \circ T_\alpha^k \right)^2 \right].$$

The full result of [22] is more general than this. It applies to the Young tower setting of [52], and for maps in the intermittent family it gives an estimate on the rate of convergence also when $1/3 \le \alpha < 1/2$, depending on the local behavior of $f(x)$ around the neutral fixed point $x = 0$.

It follows directly from the preceding paragraph using the Cramér-Wold theorem that the multivariate CLT holds for all sufficiently regular functions $f : [0, 1] \to \mathbb{R}^d$ with $d \ge 1$. However, bounds on the rate of convergence do not directly pass to higher dimensions. To estimate the speed of convergence in higher dimensions, a general approach based on Rio's method of normal approximation [44] was developed by Pène in [40]. Theorem 2.1 below is a special case of Pène's result applied to dynamical systems.

[6] Strictly speaking, it suffices to require that g is Lipschitz continuous except across a single point $z \in S^1$, when the distance of two points $x, y \in S^1 \setminus \{z\}$ is understood as the length of the arc between the points not containing $z$.

Let $(T, X, \mathcal{B}, \mu)$ be a probability preserving transformation. The covariance between any two bounded measurable functions $u, v : X \to \mathbb{R}$ is defined by

$$\mathrm{Cov}_\mu(u, v) = \mu(uv) - \mu(u)\mu(v).$$

Given a measurable function $f : X \to \mathbb{R}^d$ with $d \geq 1$, we write

$$f^k = f \circ T^k$$

for all $k \geq 0$, and denote the coordinate functions of $f^k$ by $f^k_\alpha$, $\alpha \in \{1, \dots, d\}$. We set

$$S_N = \sum_{k=0}^{N-1} f^k$$

and $W_N = N^{-1/2} S_N$ for all $N \geq 1$. Expectation of a function $h : \mathbb{R}^d \to \mathbb{R}$ with respect to the $d$-dimensional centered normal distribution $\mathcal{N}(0, \Sigma)$ with covariance matrix $\Sigma \in \mathbb{R}^{d \times d}$ is denoted by $\Phi_\Sigma(h)$, i.e.

$$\Phi_\Sigma(h) = \frac{1}{\sqrt{(2\pi)^d \det \Sigma}} \int_{\mathbb{R}^d} e^{-\frac{1}{2} w \cdot \Sigma^{-1} w} h(w)\, dw.$$

For a function $G : \mathbb{R}^d \to \mathbb{R}$, we write $D^k G$ for the $k$th derivative of $G$, and also denote $\nabla G = D^1 B$. We define

$$\|D^k G\|_\infty = \max\{\|\partial_1^{t_1} \cdots \partial_d^{t_d} G\|_\infty : t_1 + \cdots + t_d = k\}.$$

Finally, given two vectors $v, w \in \mathbb{R}^d$, we write $v \otimes w$ for the $d \times d$ matrix with entries

$$(v \otimes w)_{\alpha\beta} = v_\alpha w_\beta.$$

**Theorem 2.1.** *Let $f : X \to \mathbb{R}^d$ be a bounded measurable function with $\mu(f) = 0$. Suppose that there exist $r \in \mathbb{Z}_+$, $C \geq 1$, $M \geq \max\{1, \|f\|_\infty\}$ and a sequence of non-negative real numbers $(\varphi_{p,l})_{p,l}$ such that the following conditions hold:*

**(P1)** $\varphi_{p,l} \leq 1$ *and* $\sum_{p=1}^\infty p \max_{0 \leq l \leq \lfloor p/(r+1) \rfloor} \varphi_{p,l} < \infty$.

**(P2)** *For any integers $a, b, c$ satisfying $1 \leq a + b + c \leq 3$; for any integers $i, j, k, p, q, l$ with $0 \leq i \leq j \leq k \leq k + p \leq k + p + q \leq k + p + l$; for any $\alpha, \beta, \gamma \in \{1, \dots, d\}$; and for any bounded differentiable function $G : \mathbb{R}^d \times ([-M, M]^d)^3 \to \mathbb{R}$ with bounded gradient,*

$$\begin{aligned}&|\mathrm{Cov}_\mu[G(S_i, f^i, f^j, f^k), (f^{k+p}_\alpha)^a (f^{k+p+q}_\beta)^b (f^{k+p+l}_\gamma)^c]| \\ &\leq C(\|G\|_\infty + \|\nabla G\|_\infty)\, \varphi_{p,l}.\end{aligned}$$

*Then the limit*

$$\Sigma = \lim_{N \to \infty} \mu(W_N \otimes W_N)$$

*exists. If $\Sigma = 0$, then the sequence $(S_N)_{N \geq 0}$ is bounded in $L^2(\mu)$. Otherwise there exists $B > 0$ such that for any Lipschitz continuous function $h : \mathbb{R}^d \to \mathbb{R}$,*

$$|\mu(h(W_N)) - \Phi_\Sigma(h)| \leq \frac{B\, \mathrm{Lip}(h)}{\sqrt{N}}$$

*for all $N \geq 1$.*

The theorem guarantees the optimal rate of convergence with respect to the Kantorovich distance (Lipschitz continuous test functions) for systems satisfying conditions (P1) and (P2). Note that (P2) can be viewed as a generalized correlation decay condition where the composed outer function is not a product of one-dimensional observables but instead depends on a finite fragment of the system's trajectory. Condition (P1) requires that the decay rate, given by the numbers $\varphi_{p,l}$, should be sufficiently rapid. Pène showed in [40] that the conditions hold for the Sinai billard and Knudsen gas models, with $\varphi_{p,l}$ decaying exponentially in $p$. In article (II), we show by extending the probabilistic approximation method of [1,32] that the Pomeau-Manneville system $(T_\alpha, [0,1], \mathcal{B}([0,1]), \hat{\mu}_\alpha)$ satisfies (P2) with $\varphi_{p,l} = p^{1-1/\alpha}(\log p)^{1/\alpha}$, whenever $f : [0,1] \to \mathbb{R}^d$ is Lipschitz continuous. The conclusion of Theorem 2.1 then follows under the assumption $\alpha < 1/3$.

The primary aim of article (II) is not to extend the stationary CLT, but rather to develop tools for analyzing non-stationary systems, in particular quasistatic and sequential systems. A general operator theoretic approach to showing the CLT for sequences $(T_n)_{n=1}^\infty$ of transformations $T_n : X \to X$ was developed by Conze and Raugi in [13]. Their results roughly apply to a class of piecewise smooth uniformly expanding maps whose quasi-compact transfer operators satisfy a minoration property together with an exponential memory loss in the bounded variation norm. For such systems they proved a CLT of the following type: given a regular function $f : X \to \mathbb{R}$, and an initial distribution $\mu$ for $x \in X$, the centered Birkhoff sums

$$S_N = \sum_{k=0}^{N-1} [f \circ T_k \circ \cdots T_1 - \mu(f \circ T_k \circ \cdots T_1)]$$

satisfy

$$\frac{S_N}{\sqrt{\mathrm{Var}(S_N)}} \xrightarrow{d} \mathcal{N}(0,1), \tag{12}$$

given that $\mathrm{Var}(S_N) \to \infty$. The result was applied in a concrete model of $\beta$ transformations $T_{\beta_n}$ where the parameters $\beta_n$ approximate some fixed $\beta$ with growing $n$; in this case $\mathrm{Var}(S_N)$ grows linearly if $f$ is not a coboundary for $T_\beta$[7]. Recently, Nicol et al. [38] extended the method of [13] to a setting of intermittent maps, despite the fact that there is no spectral gap for the transfer operators. Their result established (12) for Pomeau-Manneville maps $T_n = T_{\alpha_n}$ of a suitable parameter range, provided that the growth of $\mathrm{Var}(S_N)$ is sufficiently rapid. Below we discuss quite a different approach to normal approximation, which is suited also for studying the rate of convergence in non-stationary CLTs such as (12).

There have been other recent results on limit laws for non-stationary systems, some of which extend and improve those mentioned above. For the expanding model of Conze and Raugi (and related models), almost sure invariance principles [25] and concentration inequalities [2] have been shown. In [28], the authors introduced a new version of Gordin's martingale approximation method that allowed them to examine a variety of limit laws (e.g. invariance principles) for sequences of non-uniformly hyperbolic systems. The existence of extreme value laws for sequences of Pomeau-Manneville maps was shown in [16]. For recent advances in the research of slowly mixing random dynamical systems we refer the reader to [4–7].

[7] i.e. can not be written as $g - g \circ T_\beta$

2.1. **Stein's method.** In [47], Stein introduced a method suitable for the normal approximation of weakly dependent random variables. The method has seen extensive development in the literature of probability theory (see e.g. [12, 18, 20, 43]), but a systematic adaptation for dynamical systems had not been done until recently in [26]. Like the Pène-Rio method [40], Stein's method allows to turn the problem of normal approximation into a set of correlation decay conditions. However, unlike the method of [40] or, say, the martingale CLT, Stein's method does not resort to the use of characteristic functions of the distributions. Instead, the method is based on solving the so called Stein equation

$$\operatorname{tr}\Sigma D^2 A(w) - w \cdot \nabla A(w) = h(w) - \Phi_\Sigma(h), \tag{13}$$

where $\operatorname{tr}\Sigma D^2 A(w)$ denotes the trace of the matrix $\Sigma D^2 A(w)$. Stein considered such equations (in the one-dimensional setting) with the following idea: suppose that, for each test function $h$ belonging to some class $\mathscr{H}$, the equation (13) has a solution $A$ belonging to another class of functions $\mathscr{A}$. Then,

$$\sup_{h\in\mathscr{H}} |\mathbb{E}h(W) - \Phi_\Sigma(h)| \le \sup_{A\in\mathscr{A}} |\mathbb{E}[\operatorname{tr}\Sigma D^2 A(W) - W\cdot\nabla A(W)]|. \tag{14}$$

Hence, the distance between the distributions of $W$ and $\mathcal{N}(0,\Sigma)$ can be bounded if the right side can be bounded, and this only involves working with the distribution of $W$. It is known [10, 18–20] that $\max_{1\le k\le 3}\|D^k h\|_\infty < \infty$ implies $A \in C^3(\mathbb{R}^d,\mathbb{R})$ and $\|D^k A\|_\infty < \infty$ for $1 \le k \le 3$.

In the setting of a measure preserving transformation $(T, X, \mathcal{B}, \mu)$ the aim is to again estimate the distribution of $W = N^{-1/2}\sum_{k=0}^{N-1} f^k$ by a normal distribution, where $f : X \to \mathbb{R}^d$ is $d$-dimensional and the notation is as explained above Theorem 2.1. If the system is sufficiently mixing, the right hand side of (14) can be bounded by Taylor expanding $\nabla A(W)$ about the auxiliary random vectors

$$W^n = W^n(N,K) = W - \frac{1}{\sqrt{N}}\sum_{k\in[n]_K} f^k,$$

where

$$[n]_K = [n]_K(N) = \{k \in \mathbb{Z}_+ \cap [0, N-1] \,:\, |k-n| \le K\}.$$

The random vector $W^n$ differs from $W$ by a time gap within $[0, N-1]$ of radius $K = K(N)$, centered at time $n$. In this step it is crucial that $A$ is sufficiently smooth and has bounded partial derivatives. The resulting estimate on the right side of (14), given by the following result from [26], depends on the size of the gap $K$:

**Theorem 2.2.** *Let $f : X \to \mathbb{R}^d$ be a bounded measurable function with $\mu(f) = 0$. Let $h : \mathbb{R}^d \to \mathbb{R}$ be any three times differentiable function with $\|D^k h\|_\infty < \infty$ for $1 \le k \le 3$. Fix integers $N > 0$ and $0 \le K < N$. Suppose that the following conditions are satisfied:*

- **(S1)** *There exist constants $C_2 > 0$ and $C_4 > 0$, and a non-increasing function $\varphi : \mathbb{Z}_+ \to \mathbb{R}_+$ with $\varphi(0) = 1$ and $\sum_{i=1}^\infty i\varphi(i) < \infty$, such that*

$$|\mu(f_\alpha f_\beta^k)| \le C_2\,\varphi(k)$$

$$|\mu(f_\alpha f_\beta^l f_\gamma^m f_\delta^n)| \le C_4 \min\{\varphi(l), \varphi(n-m)\}$$

$$|\mu(f_\alpha f_\beta^l f_\gamma^m f_\delta^n) - \mu(f_\alpha f_\beta^l)\mu(f_\gamma^m f_\delta^n)| \le C_4\,\varphi(m-l)$$

*hold whenever $k \ge 0$; $0 \le l \le m \le n < N$; $\alpha, \beta, \gamma, \delta \in \{\alpha', \beta'\}$ and $\alpha', \beta' \in \{1, \dots, d\}$.*

**(S2)** *There exists a function* $\tilde{\varphi} : \mathbb{Z}_+ \to \mathbb{R}_+$ *such that*

$$|\mu(f^n \cdot \nabla h(v + W^n t))| \le \tilde{\varphi}(K)$$

*holds for all* $0 \le n < N$, $0 \le t \le 1$ *and* $v \in \mathbb{R}^d$.

**(S3)** $f$ *is not a coboundary in any direction.*[8]

*Then*

$$\Sigma = \mu(f \otimes f) + \sum_{n=1}^{\infty} (\mu(f^n \otimes f) + \mu(f \otimes f^n)) \tag{15}$$

*is a well-defined, symmetric, positive-definite,* $d \times d$ *matrix; and*

$$|\mu(h(W)) - \Phi_\Sigma(h)| \le C_* \left( \frac{K+1}{\sqrt{N}} + \sum_{i=K+1}^{\infty} \varphi(i) \right) + \sqrt{N} \tilde{\varphi}(K), \tag{16}$$

*where*

$$C_* = 12 d^3 \max\{C_2, \sqrt{C_4}\} \left( \|D^2 h\|_\infty + \|f\|_\infty \|D^3 h\|_\infty \right) \sum_{i=0}^{\infty} (i+1)\varphi(i)$$

*is independent of* $N$ *and* $K$.

The result gives an estimate on the distance to a $d$-dimensional normal distribution with respect to a smooth metric. In the case $d = 1$, the smoothness of $h$ can be relaxed to Lipschitz continuity, but this comes with two expenses. First, a bound of the form (S2) must be verified for a whole class of less smooth test functions $h$. Secondly, the upper bound in the conclusion depends on $\Sigma$: even in the case of independent random variables the bound on the Kantorovich (or Wasserstein) distance depends on how the variance compares with higher moments. The smooth metric is insensitive to the size of the limit variance, as is seen from the case $d = 1$ of the above theorem. We refer the reader to [26] for details of the alternative one-dimensional formulation.

Note that the assumptions as well as the conclusion of the result concern only the given functions $f$ and $h$, and the fixed numbers $N$ and $K$. Moreover, the constant $C_*$ in the bound is expressed entirely in terms of the quantities appearing in the assumptions. Condition (S1) requires decay of correlations of orders two and four, at a rate which has a finite first moment. Condition (S2) is similar to condition (P2) of Theorem 2.1, for it requires that the components of $f^n$ and $\nabla h(v + W^n t)$ are nearly uncorrelated. Of course, for the bound in (16) to be of any use, we need $K \ll \sqrt{N}$, and $\tilde{\varphi}(K)$ needs to be small.

A remarkable feature of Stein's method is the fact that it is sufficiently flexible to be used for normal approximation in non-stationary settings. A generalization of Theorem 2.2 for sequential compositions $T_k \circ \cdots \circ T_1$ has been obtained by Olli Hella and will be reported as a part of his doctoral thesis. The result can be used, for instance, to estimate the rate of convergence in CLTs of the form (12) or in CLTs for quasistatic systems. The conclusion of the generalized result concerns the normal approximation

[8] Given a unit vector $v \in \mathbb{R}^d$, we say that $f$ is a coboundary in the direction $v$ if there exists a function $g_v : X \to \mathbb{R}$ in $L^2(\mu)$ such that

$$v \cdot f = g_v - g_v \circ T.$$

of

$$W = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \bar{f}^k,$$

where

$$\bar{f}^k = f \circ T_k \circ \cdots \circ T_1 - \mu(f \circ T_k \circ \cdots \circ T_1)$$

and nothing is assumed on the invariance of $\mu$. The result always gives an estimate on the distance to $\mathcal{N}(0, \Sigma_N)$, where $\Sigma_N = \mu(W_N \otimes W_N)$, but here a simple condition such as (S3) no longer guarantees convergence to a fixed normal distribution. For such a result one has to typically know more about the behavior of $\mu(W_N \otimes W_N)$ as $N \to \infty$ (for more discussions related to these issues see [13, 37, 38]). In addition to the above changes, one has to also recast conditions (S1) and (S2) for the sequential system. E.g., in (S2) one needs to replace the occurences of $f \circ T^k$ throughout by the function $\bar{f}^k$ defined above.

Conditions (S1)-(S3) of Theorem 2.2 were verified in [26] for Sinai billiards with convex scatterers. In this case (S1) and (S2) hold with exponentially decaying bounds, and by choosing $K \sim \log N$ the theorem then yields the nearly optimal rate $(\log N)N^{-1/2}$. In article (II) we establish (S1)-(S3) in the case of a single Pomeau-Manneville map $T_\alpha$ with $\alpha < 1/3$. For Lipschitz continuous observables $f : [0,1] \to \mathbb{R}^d$, the application produces the bound

$$\begin{aligned} &\left| \hat{\mu}_\alpha \left[ h\left( \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} f \circ T_\alpha^k \right) \right] - \Phi_\Sigma(h) \right| \\ &\le Cd^3 \max\{1, \|f\|_{\mathrm{Lip}}^3\}(\|\nabla h\|_\infty + \|D^2 h\|_\infty + \|D^3 h\|_\infty) N^{\alpha - \frac{1}{2}} (\log N)^{\frac{1}{\alpha}}, \end{aligned} \tag{17}$$

where $C > 0$ depends only on $T_\alpha$. The rate in the bound is not optimal, but note that, unlike Theorem 2.1, Stein's method enables us to maintain explicit control of the constants with respect to $f$ and $d$. The bound (17) extends to a setting of sequential intermittent systems through the generalization of Stein's method discussed in the previous paragraph.

2.2. **Functional correlation decay.** Given an admissible sequence $(T_n)_{n \in \mathbb{N}}$ of Pomeau-Manneville maps, we denote $\widetilde{T}_n = T_n \circ \cdots \circ T_1$. Then for any initial measure $\mu$ whose density belongs to $\mathcal{C}_*$, the pair correlation bound (4) tells us that

$$\left| \int F(x, \widetilde{T}_n x)\, d\mu(x) - \iint F(x, \widetilde{T}_n y)\, d\mu(x)\, d\mu(y) \right| \lesssim_F \rho(n), \tag{18}$$

whenever $F(x,y) = f(x)g(y)$ is a product of two one-dimensional observables. However, a number of limit theorems, including the CLTs mentioned above and Gouëzel's almost sure invariance principle [23], require bounds on integrals of the form $\int F \circ (\widetilde{T}_n)_{k=1}^n \, d\mu$ where $F \circ (\widetilde{T}_n)_{k=1}^n$ is a more general functional of the process. Motivated by this observation we generalize the bound (18) in article (II) as follows.

Given a function $F : [0,1]^d \to \mathbb{R}$ and $\alpha \in \{1, \ldots, d\}$, we denote by $\mathrm{Lip}(F; \alpha)$ the quantity

$$\sup_{y_1, \ldots, y_d \in [0,1]} \sup_{x_\alpha \neq \hat{x}_\alpha} \frac{|F(y_1, \ldots, y_{\alpha-1}, x_\alpha, y_{\alpha+1}, \ldots, y_d) - F(y_1, \ldots, y_{\alpha-1}, \hat{x}_\alpha, y_{\alpha+1}, \ldots, y_d)|}{|x_\alpha - \hat{x}_\alpha|},$$

and say that $F$ is Lipschitz continuous in the $\alpha$th coordinate $x_\alpha$ if $\mathrm{Lip}(F; \alpha) < \infty$.

**Theorem 2.3.** *Let $(T_n)_{n\geq 1}$ be any admissible sequence of maps. Let $F : [0,1]^{k+1} \to \mathbb{R}$ be a bounded function, and fix integers $0 \leq n_1 \leq \ldots \leq n_k$, $1 \leq l_1 < \ldots < l_p < k$. Suppose that F is Lipschitz continuous in the coordinate $x_\alpha$ whenever $1 \leq \alpha \leq l_p + 1$, and denote by $H(x_0, \ldots, x_p)$ the function*

$$F(x_0, \widetilde{T}_{n_1}(x_0), \ldots, \widetilde{T}_{n_{l_1}}(x_0), \widetilde{T}_{n_{l_1+1}}(x_1), \ldots, \widetilde{T}_{n_{l_2}}(x_1), \ldots, \widetilde{T}_{n_{l_p+1}}(x_p), \ldots, \widetilde{T}_{n_k}(x_p)).$$

*Then, for any probability measures $\mu, \mu_1, \ldots, \mu_p$ whose densities belong to $\mathcal{C}_*$,*

$$\left| \int H(x, \ldots, x)\, d\mu(x) - \int \cdots \int H(x_0, \ldots, x_p)\, d\mu(x_0) d\mu_1(x_1) \ldots d\mu_p(x_p) \right|$$
$$\leq C(\|F\|_\infty + \max_{1\leq\alpha\leq l_p+1} \mathrm{Lip}(F;\alpha)) \sum_{i=1}^{p} \rho(n_{l_i+1} - n_{l_i}),$$

*where $C > 0$ depends only on $T_{\beta_*}$, $\rho(n) = n^{-\frac{1}{\beta_*}+1}(\log n)^{\frac{1}{\beta_*}}$ for $n \geq 2$, and $\rho(0) = \rho(1) = 1$.*

In the special case $p = 1$ of a single "gap", the bound simplifies to

$$\left| \int H(x,x)\, d\mu(x) - \iint H(x,y)\, d\mu(x)\, d\mu_1(y) \right| \lesssim \rho(n_{l+1} - n_l), \tag{19}$$

where $l = l_1$ and

$$H(x,y) = F(x, \widetilde{T}_{n_1}(x), \ldots, \widetilde{T}_{n_l}(x), \widetilde{T}_{n_{l+1}}(y), \ldots, \widetilde{T}_{n_k}(y)).$$

Then, given a Lipschitz continuous observable $f : [0,1] \to \mathbb{R}^d$, a smooth test function $h : \mathbb{R}^d \to \mathbb{R}$, and $0 < K < n \leq N-1$, we might for instance apply the bound with

$$H(x,y) = F(x, T_\alpha x, \ldots, T_\alpha^{n-K} x, T_\alpha^n y, T_\alpha^{n+K} y, \ldots, T^N y),$$

where

$$F(x_0, \ldots, x_{n-K}, x_n, x_{n+K}, \ldots, x_{N-1}) = f(x_n) \cdot \nabla h \left( v + \frac{1}{\sqrt{N}} \sum_{i \notin [n]_K} f(x_i) t \right).$$

The application yields condition (S2) of Theorem 2.2 with $\tilde{\varphi}(K) = O(\rho(K))$.

The full result of Theorem 2.3 follows by induction from the special case (19). The proof of (19) is based on partitioning the unit interval $[0,1] = \cup_{\theta=1}^{2^N} I_\theta$ into suitably small subintervals $I_\theta = I_\theta(N)$ such that $\widetilde{T}_N \restriction I_\theta$ maps each $I_\theta$ diffeomorphically onto $(0,1)$. The partition enables us to utilize the continuity of $x \mapsto H(x,y)$: if $c_\theta$ denotes the midpoint of $I_\theta$, we can approximate $H(x,y) \approx H(c_\theta, y)$ with a small error. On the other hand, $H(c_\theta, y) = G(\widetilde{T}_{n_l+1} y)$ for some function $G$ with $\|G\|_\infty \leq \|F\|_\infty$. It follows that

$$\int H(x,x)\, d\mu(x) - \iint H(x,y)\, d\mu(x)\, d\mu_1(y)$$
$$= \sum_{\theta=1}^{2^{n_*}} \int_{I_\theta} \left( H(x,x) - \int H(x,y)\, d\mu_1(y) \right) d\mu(x)$$
$$\approx \sum_{\theta=1}^{2^{n_*}} \int_{I_\theta} \left( H(c_\theta,x) - \int H(c_\theta,y)\, d\mu_1(y) \right) d\mu(x),$$

where the error is small. Then to obtain (19), it remains to approximate

$$\frac{1}{\mu(I_\theta)} \int_{I_\theta} H(c_\theta, x)\, d\mu(x) \approx \int H(c_\theta, y)\, d\mu_1(y) \tag{20}$$

for each $\theta$ with a good upper bound on the error. For this, we need a version of polynomial memory loss for conditional densities of the form

$$\frac{1}{\mu(I_\theta)}\mathbf{1}_{I_\theta}\frac{dm}{d\mu}, \tag{21}$$

where $\mathbf{1}_{I_\theta}$ denotes the indicator function of $I_\theta$. We prove in article (II) the following bound:

**Proposition 2.4.** *Let $h, g \in \mathcal{C}_*$ be densities, and let $m \ge 1$ be an integer. Define*

$$h_\theta = \mu(I_\theta)^{-1}\mathbf{1}_{I_\theta}h \quad \forall \theta \in \{1, \dots, 2^N\}.$$

*Then,*

$$\sum_{\theta=1}^{2^N} \mu(I_\theta)\|\mathcal{L}_{N+m}\cdots\mathcal{L}_1(h_\theta - g)\|_1 \lesssim \rho(m).$$

The bound follows by observing that the method of [1, 32] can be extended from cone densities $h$ to conditional densities of the form $h_\theta$. The result plays an instrumental role in article (III) where we utilize similar partitions $\{I_\theta\}_{\theta=1}^{2^N}$ in the setting of quasistatic systems.

We refer the reader to [30] for functional bounds and their applications to limit theorems in the setting of Sinai billiards.

## 3. Statistical properties of intermittent QDSs

Recall that a QDS is a pair $(\mathbf{T}, \tau)$ where $\mathbf{T} = \{T_{n,k} : 0 \le k \le n, n \in \mathbb{N}\}$ is a triangular array of maps in a topological space $\mathcal{M}$, and $\tau : [0,1] \to \mathcal{M}$ is a curve such that $T_{n,\lfloor nt \rfloor} \to \tau_t$ as $n \to \infty$. In the case of the intermittent family we adapt the abstract definition as follows:

**Definition 3.1** (Intermittent QDS)**.** *Let $X = [0,1]$ and $\mathcal{M} = \{T_\alpha : 0 \le \alpha \le 1\}$ (equipped, say, with the uniform topology). Next, let*

$$\{\alpha_{n,k} \in [0,1] \,:\, 0 \le k \le n,\ n \ge 1\}$$

*be a triangular array of parameters and*

$$\gamma : [0,1] \to [0,1)$$

*a piecewise continuous curve satisfying*

$$\lim_{n\to\infty} \alpha_{n,\lfloor nt \rfloor} = \gamma_t \tag{22}$$

*for all $t$. Finally, define $\tau_t = T_{\gamma_t}$ and*

$$T_{n,k} = T_{\alpha_{n,k}}.$$

Recall that for each $\alpha \in [0,1)$ there exists an absolutely continuous $T_\alpha$-invariant probability measure $\hat{\mu}_\alpha$. We set $\hat{\nu}_t = \hat{\mu}_{\gamma_t}$ so that $\hat{\nu}_t$ is invariant for $\tau_t = T_{\gamma_t}$. The transfer operator associated to $T_{n,k}$ is denoted by $\mathcal{L}_{n,k} = \mathcal{L}_{\alpha_{n,k}}$.

Sections 3.1 and 3.2, respectively, summarize articles (I) and (III) where we show results for intermittent QDSs that extend those of [14, 49]. In the analysis we utilize extensively the known theory of sequential Pomeau-Manneville maps, in particular the polynomial memory loss result of [1], but we also develop new tools such as perturbation estimates. The single most significant difference to [14, 49] is that the QDS considered here is non-uniformly hyperbolic.

To obtain our results we need to make some assumptions regarding the regularity of $\gamma$ and the rate of convergence in (22). Throughout this section we assume that $\gamma$ is Hölder continuous of order $\eta \in (0,1]$, that

$$\overline{\gamma([0,1])} \subset [0, \beta_*] \tag{23}$$

holds for some $\beta_* \in (0,1)$, and that

$$\lim_{n\to\infty} n^{\eta} \sup_{t\in[0,1]} |\alpha_{n,\lfloor nt \rfloor} - \gamma_t| < \infty. \tag{24}$$

Condition (23) enables us to maintain uniform control in estimates involving memory loss, whereas condition (24) reflects the regularity of $\gamma$. For simplicity we do not allow $\gamma$ to have jumps here, although the results in article (I) are proven for piecewise Hölder continuous $\gamma$ (also the results of article (III) extend to such settings).

3.1. **Ergodic properties.** For clarity we recast some of the definitions from Section 1.2 for the intermittent QDS: given a bounded measurable function $f : [0,1] \to \mathbb{R}$, we denote

$$f_{n,k} = f \circ T_{n,k} \circ \cdots \circ T_{n,1}, \quad 0 \le k \le n,$$

and define the functions $S_n : [0,1] \times [0,1] \to \mathbb{R}$ by

$$S_n(x,t) = \int_0^{nt} f_{n,\lfloor s \rfloor}(x)\, ds, \quad n \ge 1.$$

Then, given an initial measure $\mu$ with density in $\mathcal{C}_*$, we set

$$\zeta_n(x,t) = n^{-1} S_n(x,t) = \int_0^t f_{n,\lfloor ns \rfloor}(x)\, ds,$$

and

$$\bar{\zeta}_n(x,t) = \int_0^t \bar{f}_{n,\lfloor ns \rfloor}(x)\, ds,$$

where

$$\bar{f}_{n,k} = f - \mu(f_{n,k}).$$

Finally, we define

$$\zeta(t) = \int_0^t \hat{\mu}_{\gamma_s}(f)\, ds.$$

The purpose of article (I) is to extend the ergodic theorems of [49] by showing that

$$\lim_{n\to\infty} \sup_{t\in[0,1]} |\zeta_n(x,t) - \zeta(t)| = 0 \tag{25}$$

for almost every $x \in [0,1]$ with respect to Lebesgue measure, whenever $f \in C^0([0,1])$. By a density argument we see that it suffices to prove the result for all $f \in C^1([0,1])$. As explained under Theorem 1.5 in Section 1.2, (25) is a consequence of

$$\lim_{n\to\infty} \sup_{t\in[0,1]} \left| \int_0^t \mu(f_{n,\lfloor ns \rfloor}) - \hat{\mu}_{\gamma_s}(f)\, ds \right| = 0, \tag{26}$$

provided that the following condition holds: for all integers $2 \le l \le 4$, $j \in \{1, l-1\}$ and $0 \le k_1 \le \cdots \le k_l$,

$$|\mu(f_{n,k_1} \cdots f_{n,k_l}) - \mu(f_{n,k_1} \cdots f_{n,k_j}) \mu(f_{n,k_{j+1}} \cdots f_{n,k_l})| \lesssim_f \Phi(k_{j+1} - k_j), \tag{27}$$

where $\Phi(s) = s^{-1}(\log s)^{-2}$ if $s \geq 2$, and $\Phi(s) = 2^{-1}(\log 2)^{-2}$ if $0 < s < 2$. For this reason we prove in article (I) that if $s \neq 0$, then $\mu(f_{n,\lfloor ns \rfloor})$ approximates $\hat{\mu}_{\gamma_s}(f)$ with increasing $n$ as follows:

**Proposition 3.2.** *Let $\mu$ be a probability measure with density $h \in \mathcal{C}_*$, and let $\mu_{n,k} = (T_{n,k} \circ \cdots T_{n,1})_*\mu$ be its pushforward with density $h_{n,k} = \mathcal{L}_{n,k} \cdots \mathcal{L}_{n,1}h$. There exist $c_1 > 0$ and $p_0, p_1 \in (0,1)$ such that*

$$\|h_{n,k} - \hat{h}_{\alpha_{n,k}}\|_{L^1(m)} \lesssim n^{-p_0}$$

*whenever $c_1 n^{p_1} < k \leq n$. Moreover, given a bounded function $f : [0,1] \to \mathbb{R}$,*

$$|\mu_{n,\lfloor nt \rfloor}(f) - \hat{\mu}_{\gamma_t}(f)| \lesssim \|f\|_\infty n^{-p_0} \tag{28}$$

*whenever $c_1 n^{p_1 - 1} < t \leq 1$.*

Since $f$ is bounded, (26) follows at once from (28), meaning that we obtain (25) under condition (27). Condition (27) applies only if $\beta_*$ is sufficiently small: by a correlation decay result, such as Theorem 4.1 of article (I) or the functional bound discussed in the previous section, we have that

$$|\mu(f_{n,k_1} \cdots f_{n,k_l}) - \mu(f_{n,k_1} \cdots f_{n,k_j})\mu(f_{n,k_{j+1}} \cdots f_{n,k_l})| \lesssim_f \rho(k_{j+1} - k_j),$$

where $\rho(k) \lesssim \Phi(k)$ iff $\beta_* < 1/2$.

The proof of Proposition 3.2 requires us to control $\|(\mathcal{L}_{n,k} - \mathcal{L}_{n,j})h\|_{L^1(m)}$ for densities $h \in \mathcal{C}_*$. For this we show in article (I) the following $L^1$-perturbation estimate:

**Theorem 3.3.** *Let $0 < \beta_* < 1$. Then,*

$$\|(\mathcal{L}_\alpha - \mathcal{L}_\beta)h\|_{L^1(m)} \lesssim \|h\|_{L^1(m)}(\beta - \alpha)^{\frac{1}{3}(1-\beta_*)}|\log(\beta - \alpha)| \qquad \forall\, h \in \mathcal{C}_* \tag{29}$$

*and*

$$\|\hat{h}_\alpha - \hat{h}_\beta\|_{L^1(m)} \lesssim (\beta - \alpha)^{\frac{1}{3}(1-\beta_*)^2}|\log(\beta - \alpha)|^{\frac{1}{\beta_*}} \tag{30}$$

*hold whenever $0 \leq \alpha < \beta \leq \beta_*$.*

The result is instrumental in article (III). It applies also to more general functions $h$ through the following observation which is (essentially) from [32]: given any admissible sequence $(T_n)_{n\geq 1}$, Lipschitz continuous functions $f_1, f_2$, $h \in \mathcal{C}_*$ and $n \geq 0$, there exist $g_1, \ldots, g_4 \in \mathcal{C}_*$ such that

$$f_2 \cdot \mathcal{L}_n \cdots \mathcal{L}_1(f_1 h) = g_1 - g_2 + g_3 - g_4$$

and

$$\|g_i\|_{L^1(m)} \lesssim \|f_1\|_{\mathrm{Lip}}\|f_2\|_{\mathrm{Lip}} m(h).$$

We provide a detailed proof for this statement in articles (I) and (III).

Above we have discussed how one can prove an ergodic theorem in the parameter range $\beta_* < 1/2$. In the parameter range $\beta_* \in [1/2, 1)$, the objective is still to show that $\lim_{n\to\infty} \bar{\zeta}_n(x,t) = 0$ almost surely, but here the fourth moment condition breaks down and we can no longer rely on the strategy of [49]. Instead we invoke the following known refinement of the Cauchy condensation criterion [15,33]:

**Proposition 3.4.** *Let the numbers $a_n \geq 0$, $n \geq 1$, satisfy*

$$\sum_{n=1}^{\infty} \frac{a_n}{n} < \infty.$$

*Then there exists an increasing sequence* $(n_k)_{k\geq 1}$ *such that*

$$\sum_{k=1}^{\infty} a_{n_k} < \infty \quad \text{and} \quad \lim_{k\to\infty} \frac{n_{k+1}}{n_k} = 1.$$

We verify in article (I) that the result implemented with

$$a_n = \mu \left[ \left( \sup_{0\leq k\leq \lfloor \log n\rfloor} \left| \bar{\zeta}_n \left( \tfrac{k}{\lfloor \log n\rfloor} \right) \right| \right)^2 \right]$$

implies the almost sure limit

$$\lim_{k\to\infty} \sup_{t\in[0,1]} |\bar{\zeta}_{n_k}(x,t)| = 0 \tag{31}$$

for a subsequence $(n_k)$ such that $n_k/n_{k+1} \to 1$. The argument can be repeated for any subsequence $(n_k)$ to obtain a further subsequence $(n_{k_l})$ that satisfies (31). That is, $\sup_{t\in[0,1]} |\bar{\zeta}_n(t)| \to 0$ in probability. We are unable to strengthen (31) to almost sure convergence of the full sequence in the wider parameter range. The issues that arise with methods known to us have to do with the necessity of controlling

$$\zeta_n(t) - \zeta_m(t) = \int_0^t f \circ T_{n,\lfloor ns\rfloor} \circ \cdots \circ T_{n,1} - f \circ T_{m,\lfloor ms\rfloor} \circ \cdots \circ T_{m,1}\, ds$$

with $n \neq m$, which involves comparing different levels of the array $\mathbf{T}$, consisting of different maps.

To summarize the above discussion, we state as a result the conclusions of article (I) concerning almost sure convergence.

**Theorem 3.5.** *Suppose that the curve* $\gamma : [0,1] \to [0,1]$ *is (piecewise) Hölder continuous of order* $\eta \in (0,1]$, *that*

$$\overline{\gamma([0,1])} \subset [0,\beta_*]$$

*for some* $\beta_* \in (0,1)$, *and that*

$$\lim_{n\to\infty} n^{\eta} \sup_{t\in[0,1]} |\alpha_{n,\lfloor nt\rfloor} - \gamma_t| < \infty.$$

(i) *If* $\beta_* \geq \frac{1}{2}$, *then for each* $f \in C^0([0,1])$,

$$\lim_{n\to\infty} \sup_{t\in[0,1]} |\zeta_n(x,t) - \zeta(t)| = 0 \tag{32}$$

*in probability, with respect to the Lebesgue measure. That is,*

$$\lim_{n\to\infty} m\left( \sup_{t\in[0,1]} |\zeta_n(x,t) - \zeta(t)| \geq \varepsilon \right) = 0$$

*for all* $\varepsilon > 0$.

(ii) *If* $\beta_* < \frac{1}{2}$, *then* (32) *holds for almost every* $x \in [0,1]$ *with respect to the Lebesgue measure.*

3.2. **Distributional properties.** Let $\mu$ be an initial distribution of $x \in [0,1]$ with density $h \in \mathcal{C}_*$, and let $f : [0,1] \to \mathbb{R}$ be a Lipschitz continuous function. For each $t \in [0,1]$ we denote $\hat{f}_t = f - \hat{\mu}_{\gamma_t}(f)$. In article (III), we obtain results that extend those of [14] by showing that the fluctuations

$$\xi_n(x,t) = n^{-\frac{1}{2}} S_n(x,t) - n^{-\frac{1}{2}} \mu(S_n(\cdot,t))$$

converge weakly to the law of the diffusion process

$$\xi(t) = \int_0^t \hat{\sigma}_s(f)\, dW_s,$$

where $W$ denotes a standard Brownian motion, the stochastic integral is defined in the sense of Itō, and

$$\hat{\sigma}_t^2(f) = \lim_{m\to\infty} \hat{\mu}_{\gamma_t}\left[\left(\frac{1}{\sqrt{m}} \sum_{k=0}^{m-1} \hat{f}_t \circ T_{\gamma_t}^k\right)^2\right].$$

Note that, under condition (23), the latter limit exists when $\beta_* < 1/2$.

We approach the problem by following the general strategy used in [14], outlined below Theorem 1.6. A successful implementation of the strategy requires us to verify the following two conditions:

(T) The distributions $\mathbb{P}_n^\mu$ of $\xi_n$ form a tight sequence of measures (with respect to $\mu$).

(M) The weak limit of a subsequence $(\mathbb{P}_{n_k}^\mu)$ satisfies the martingale property of Proposition 1.8.

We establish condition (M) in the full parameter range $\beta_* < 1/2$ but manage to show condition (T) only if $\beta_* < 1/3$: in the smaller parameter range condition (T) follows by the Kolmogorov-Chentsov criterion combined with polynomial correlation decay. Consequently, in the wider parameter range $1/3 \le \beta_* < 1/2$ we obtain the diffusive limit only under the assumption that $(\mathbb{P}_n^\mu)$ is tight.

The proof of condition (M) uses Taylor expansion. The role of the following estimates on second moments is to facilitate controlling the error terms in the expansion.

**Proposition 3.6.** *For all* $0 \le t \le t+\delta \le 1$,

$$\mu[[\xi_n(t+\delta) - \xi_n(t)]^2] = \int_t^{t+\delta} \hat{\sigma}_s^2(f)\, ds + \delta o(1) + o(n^{-\frac{1}{2}}),$$

*where the error terms are uniform in* $t$ *and* $\delta$.

**Proposition 3.7.** *For all* $A \in C^\infty(\mathbb{R})$, $0 \le s \le t \le 1$,

$$\mu[A(\xi_n(s))[\xi_n(t) - \xi_n(s)]^2] - \mu[A(\xi_n(s)]\mu[[\xi_n(t) - \xi_n(s)]^2] = o(1),$$

*where the error term is uniform in* $t$ *and* $s$.

The former result yields a rate at which the second moments converge toward the limiting variance, while the latter result is a type of decorrelation estimate for $\xi_n$. Key elements in the proof of Proposition 3.6 are polynomial memory loss (Theorem 1.1), perturbation of transfer operators (Theorem 3.3), and convergence of the pushforward measures $(T_{n,\lfloor nt\rfloor} \circ \cdots T_{n,1})_*\mu$ toward the SRB measure $\hat{\mu}_{\gamma_t}$ (Proposition 3.2).

The proof of Proposition 3.7, like the proof of the functional bound discussed in Section 2.2, is based on a canonical partition of the unit interval: given $0 \le s < t < 1$, we write the unit interval $[0,1] = \cup_{\theta=1}^{2^N} I_\theta$ as a union of $2^N$ subintervals, such that for each $\theta$ the map $(T_{n,N} \circ \cdots \circ T_{n,1}) \restriction I_\theta$ is a diffeomorphism onto $(0,1)$. If $N = N(n,s) \leq n$ is suitably chosen, then $x \mapsto \xi_n(x,s)$ is almost constant on each partition element $I_\theta$, and the proof of Proposition 3.7 reduces to showing

$$\sum_{\theta=1}^{2^N} \mu(I_\theta) |\mu_\theta[[\xi_n(t) - \xi_n(s)]^2] - \mu[[\xi_n(t) - \xi_n(s)]^2]| = o(1), \tag{33}$$

where $\mu_\theta$ denotes the measure $\mu$ conditioned to the interval $I_\theta$. It then remains to replace the conditional measures $\mu_\theta$ with $\mu$ in the double integral

$$\mu_\theta[[\xi_n(t) - \xi_n(s)]^2] = n \iint_{[s,t]\times[s,t]} \mu_\theta(\bar{f}_{n,\lfloor nu \rfloor} \bar{f}_{n,\lfloor nv \rfloor})\, du\, dv,$$

and control the error. The latter task is facilitated by removing a small region $[s, s + n^{-p}] \times [s, s + n^{-p}]$ from the domain of integration $[s,t] \times [s,t]$, for then we can implement Proposition 2.4 to control the remaining term. This step requires rather careful estimation using the inequalities of Jensen and Cauchy-Schwarz in combination with Proposition 2.4, but the procedure eventually leads to (33). A technical issue here is that, unlike in the model of [14], we have no uniform control on how the conditional measures $\mu_\theta$ evolve with the dynamic.

We end the discussion by stating the main result of article (III), which instead of $\xi_n$ concerns more general processes of the form

$$\chi_n^\nu(x,t) = n^{-\frac{1}{2}} S_n(x,t) - n^{-\frac{1}{2}} \nu(S_n(\cdot,t)),$$

where $\nu$ is a probability measure possibly different from the initial measure $\mu$. The generalization is obtained from the result for $\xi_n$ by an application of the Portmanteau theorem. Let us denote by $\mathbb{P}_n^{\mu,\nu}$ the distribution of $x \mapsto \chi_n^\nu(x,\cdot)$ with respect to $\mu$.

**Theorem 3.8.** *Let $f : [0,1] \to \mathbb{R}$ be Lipschitz continuous, and let the initial measure $\mu$ be such that its density belongs to $\mathcal{C}_*$. Suppose that $\gamma : [0,1] \to [0,1)$ is Hölder-continuous of order $\eta \in (0,1]$, that $\gamma([0,1]) \subset [0,\beta_*]$ for some $\beta_* < \frac{1}{2}$, and that*

$$\lim_{n\to\infty} n^\eta \sup_{t\in[0,1]} |\alpha_{n,\lfloor nt \rfloor} - \gamma_t| < \infty.$$

*If the sequence of measures $(\mathbb{P}_n^\mu)_{n\geq 1}$ is tight, then for any probability measure $\nu$, whose density $g = g_1 - g_2$ for some $g_1, g_2 \in \mathcal{C}_*$, the sequence $(\mathbb{P}_n^{\mu,\nu})_{n\geq 1}$ converges weakly to the law of the process*

$$\chi(t) = \int_0^t \hat{\sigma}_s(f)\, dW_s.$$

*Moreover, if $\beta_* < 1/3$, then $(\mathbb{P}_n^\mu)_{n\geq 1}$ is tight.*

## References


[1] Romain Aimino, Huyi Hu, Matthew Nicol, Andrei Török, and Sandro Vaienti. Polynomial loss of memory for maps of the interval with a neutral fixed point. *Discrete Contin. Dyn. Syst.*, 35(3):793–806, 2015. URL: http://dx.doi.org/10.3934/dcds.2015.35.793.

[2] Romain Aimino and Jérôme Rousseau. Concentration inequalities for sequential dynamical systems of the unit interval. *Ergodic Theory Dynam. Systems*, 36(8):2384–2407, 2016. URL: http://dx.doi.org/10.1017/etds.2015.19.

[3] José F. Alves, Stefano Luzzatto, and Vilton Pinheiro. Markov structures and decay of correlations for non-uniformly expanding dynamical systems. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 22(6):817–839, 2005. URL: https://doi.org/10.1016/j.anihpc.2004.12.002.

[4] Wael Bahsoun and Christopher Bose. Mixing rates and limit theorems for random intermittent maps. *Nonlinearity*, 29(4):1417–1433, 2016. URL: http://dx.doi.org/10.1088/0951-7715/29/4/1417.

[5] Wael Bahsoun, Christopher Bose, and Yuejiao Duan. Decay of correlation for random intermittent maps. *Nonlinearity*, 27(7):1543–1554, 2014. URL: http://dx.doi.org/10.1088/0951-7715/27/7/1543.

[6] Wael Bahsoun, Christopher Bose, and Marks Ruziboev. Quenched decay of correlations for slowly mixing systems. 2017. Preprint. arXiv:1706.04158.

[7] Wael Bahsoun, Marks Ruziboev, and Benoît Saussol. Linear response for random dynamical systems. 2017. Preprint. arXiv:1710.03706.

[8] V. I. Bakhtin. Random processes generated by a hyperbolic sequence of mappings. I. *Izv. Ross. Akad. Nauk Ser. Mat.*, 58(2):40–72, 1994. URL: https://doi.org/10.1070/IM1995v044n02ABEH001596.

[9] V. I. Bakhtin. Random processes generated by a hyperbolic sequence of mappings. II. *Izv. Ross. Akad. Nauk Ser. Mat.*, 58(3):184–195, 1994. URL: https://doi.org/10.1070/IM1995v044n03ABEH001616.

[10] Andrew Barbour. Stein's method for diffusion approximations. *Probab. Theory Related Fields*, 84(3):297–322, 1990. doi:10.1007/BF01197887.

[11] Henk Bruin, Stefano Luzzatto, and Sebastian Van Strien. Decay of correlations in one-dimensional dynamics. *Ann. Sci. École Norm. Sup. (4)*, 36(4):621–646, 2003. URL: https://doi.org/10.1016/S0012-9593(03)00025-9.

[12] Sourav Chatterjee and Elizabeth Meckes. Multivariate normal approximation using exchangeable pairs. *ALEA Lat. Am. J. Probab. Math. Stat.*, 4:257–283, 2008.

[13] Jean-Pierre Conze and Albert Raugi. Limit theorems for sequential expanding dynamical systems on $[0,1]$. In *Ergodic theory and related fields*, volume 430 of *Contemp. Math.*, pages 89–121. Amer. Math. Soc., Providence, RI, 2007. doi:10.1090/conm/430/08253.

[14] Neil Dobbs and Mikko Stenlund. Quasistatic dynamical systems. 2016. To appear in Ergodic Theory and Dynamical Systems. doi:10.1017/etds.2016.9.

[15] Aryeh Dvoretzky. On the strong stability of a sequence of events. *Ann. Math. Statistics*, 20:296–299, 1949.

[16] Ana Cristina Moreira Freitas, Jorge Milhazes, and Sandro Vaienti. Extreme value laws for sequences of intermittent maps. 2016. Preprint. arXiv:1605.06287.

[17] P. Gaspard and X.-J. Wang. Sporadicity: between periodic and chaotic dynamical behaviors. *Proc. Nat. Acad. Sci. U.S.A.*, 85(13):4591–4595, 1988. URL: https://doi.org/10.1073/pnas.85.13.4591.

[18] Robert E. Gaunt. Rates of convergence in normal approximation under moment conditions via new bounds on solutions of the Stein equation. *J. Theoret. Probab.*, 29(1):231–247, 2016. doi:10.1007/s10959-014-0562-z.

[19] Larry Goldstein and Yosef Rinott. Multivariate normal approximations by Stein's method and size bias couplings. *J. Appl. Probab.*, 33(1):1–17, 1996. doi:10.2307/3215259.

[20] Friedrich Götze. On the rate of convergence in the multivariate CLT. *Ann. Probab.*, 19(2):724–739, 1991. doi:doi:10.1214/aop/1176990448.

[21] Sébastien Gouëzel. Sharp polynomial estimates for the decay of correlations. *Israel J. Math.*, 139:29–65, 2004. URL: http://dx.doi.org/10.1007/BF02787541.

[22] Sébastien Gouëzel. Berry-Esseen theorem and local limit theorem for non uniformly expanding maps. *Ann. Inst. H. Poincaré Probab. Statist.*, 41(6):997–1024, 2005. URL: https://doi.org/10.1016/j.anihpb.2004.09.002.

[23] Sébastien Gouëzel. Almost sure invariance principle for dynamical systems by spectral methods. *Ann. Probab.*, 38(4):1639–1671, 2010. URL: https://doi.org/10.1214/10-AOP525.

[24] Chinmaya Gupta, William Ott, and Andrei Török. Memory loss for time-dependent piecewise expanding systems in higher dimension. *Math. Res. Lett.*, 20(1):141–161, 2013. URL: https://doi.org/10.4310/MRL.2013.v20.n1.a12.

[25] Nicolai Haydn, Matthew Nicol, Andrew Török, and Sandro Vaienti. Almost sure invariance principle for sequential and non-stationary dynamical systems. *Trans. Amer. Math. Soc.*, 369(8):5293–5316, 2017. doi:10.1090/tran/6812.

[26] Olli Hella, Juho Leppänen, and Mikko Stenlund. Stein's method for dynamical systems. *prerint*. URL: https://arxiv.org/abs/1701.02966.

[27] Huyi Hu. Decay of correlations for piecewise smooth maps with indifferent fixed points. *Ergodic Theory Dynam. Systems*, 24(2):495–524, 2004. URL: http://dx.doi.org/10.1017/S0143385703000671.

[28] A. Korepanov, Z. Kosloff, and I. Melbourne. Martingale–coboundary decomposition for families of dynamical systems. *Annales de l'Institut Henri Poincare (C) Non Linear Analysis*, 2017. URL: http://www.sciencedirect.com/science/article/pii/S0294144917301014, doi:https://doi.org/10.1016/j.anihpc.2017.08.005.

[29] A. Lambert, S. Siboni, and S. Vaienti. Statistical properties of a nonuniformly hyperbolic map of the interval. *J. Statist. Phys.*, 72(5-6):1305–1330, 1993. URL: https://doi.org/10.1007/BF01048188.

[30] Juho Leppänen and Mikko Stenlund. A note on the finite-dimensional distributions of dispersing billiard processes. *J. Stat. Phys.*, 168(1):128–145, 2017. URL: https://doi.org/10.1007/s10955-017-1790-2.

[31] Carlangelo Liverani. Central limit theorem for deterministic systems. In *International Conference on Dynamical Systems (Montevideo, 1995)*, volume 362 of *Pitman Res. Notes Math. Ser.*, pages 56–75. Longman, Harlow, 1996.

[32] Carlangelo Liverani, Benoît Saussol, and Sandro Vaienti. A probabilistic approach to intermittency. *Ergodic Theory Dynam. Systems*, 19(3):671–685, 1999. URL: http://dx.doi.org/10.1017/S0143385799133856.

[33] Russell Lyons. Strong laws of large numbers for weakly correlated random variables. *Michigan Math. J.*, 35(3):353–359, 1988. URL: https://doi.org/10.1307/mmj/1029003816.

[34] Nader Masmoudi and Lai-Sang Young. Ergodic theory of infinite dimensional systems with applications to dissipative parabolic PDEs. *Comm. Math. Phys.*, 227(3):461–481, 2002. URL: https://doi.org/10.1007/s002200200639.

[35] Anushaya Mohapatra and William Ott. Memory loss for nonequilibrium open dynamical systems. *Discrete Contin. Dyn. Syst.*, 34(9):3747–3759, 2014. doi:10.3934/dcds.2014.34.3747.

[36] Makoto Mori. On the intermittency of a piecewise linear map (Takahashi model). *Tokyo J. Math.*, 16(2):411–428, 1993. URL: https://doi.org/10.3836/tjm/1270128495.

[37] Péter Nándori, Domokos Szász, and Tamás Varjú. A central limit theorem for time-dependent dynamical systems. *J. Stat. Phys.*, 146(6):1213–1220, 2012. doi:10.1007/s10955-012-0451-8.

[38] Matthew Nicol, Andrew Török, and Sandro Vaienti. Central limit theorems for sequential and random intermittent dynamical systems. 2016. To appear in Ergodic Theory and Dynamical Systems. doi:10.1017/etds.2016.69.

[39] William Ott, Mikko Stenlund, and Lai-Sang Young. Memory loss for time-dependent dynamical systems. *Math. Res. Lett.*, 16(3):463–475, 2009. doi:10.4310/MRL.2009.v16.n3.a7.

[40] Françoise Pène. Rate of convergence in the multidimensional central limit theorem for stationary processes. Application to the Knudsen gas and to the Sinai billiard. *Ann. Appl. Probab.*, 15(4):2331–2392, 2005. URL: http://dx.doi.org/10.1214/105051605000000476.

[41] Giulio Pianigiani. First return map and invariant measures. *Israel J. Math.*, 35(1-2):32–48, 1980. URL: https://doi.org/10.1007/BF02760937.

[42] Yves Pomeau and Paul Manneville. Intermittent transition to turbulence in dissipative dynamical systems. *Comm. Math. Phys.*, 74(2):189–197, 1980. URL: http://projecteuclid.org/euclid.cmp/1103907981.

[43] Gesine Reinert and Adrian Röllin. Multivariate normal approximation with Stein's method of exchangeable pairs under a general linearity condition. *Ann. Probab.*, 37(6):2150–2173, 2009. URL: https://doi.org/10.1214/09-AOP467.

[44] Emmanuel Rio. Sur le théorème de Berry-Esseen pour les suites faiblement dépendantes. *Probab. Theory Related Fields*, 104(2):255–282, 1996. doi:10.1007/BF01247840.

[45] L. C. G. Rogers and David Williams. *Diffusions, Markov processes, and martingales. Vol. 2*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000. Itô calculus, Reprint of the second (1994) edition. doi:10.1017/CBO9781107590120.

[46] Omri Sarig. Subexponential decay of correlations. *Invent. Math.*, 150(3):629–653, 2002. URL: https://doi.org/10.1007/s00222-002-0248-5.

[47] Charles Stein. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. pages 583–602, 1972.

[48] Mikko Stenlund. Non-stationary compositions of Anosov diffeomorphisms. *Nonlinearity*, 24(10):2991–3018, 2011. URL: https://doi.org/10.1088/0951-7715/24/10/016.

[49] Mikko Stenlund. An almost sure ergodic theorem for quasistatic dynamical systems. *Math. Phys. Anal. Geom.*, 19(3):Art. 14, 18, 2016. doi:10.1007/s11040-016-9217-x.

[50] Mikko Stenlund, Lai-Sang Young, and Hongkun Zhang. Dispersing billiards with moving scatterers. *Comm. Math. Phys.*, 322(3):909–955, 2013. URL: https://doi.org/10.1007/s00220-013-1746-6.

[51] Maximilian Thaler. Estimates of the invariant densities of endomorphisms with indifferent fixed points. *Israel J. Math.*, 37(4):303–314, 1980. URL: https://doi.org/10.1007/BF02788928.

[52] Lai-Sang Young. Recurrence times and rates of mixing. *Israel J. Math.*, 110:153–188, 1999. URL: http://dx.doi.org/10.1007/BF02808180.